1 Lecture 1

1.1 Vector spaces

Let F be a field. Informally, a field is a set of numbers equipped with the operations $+, -, \times, \div$ that satisfy the usual arithmetic rules. A precise mathematical definition is not required in this course; for a rigorous treatment and an introduction to the beautiful Galois theory of fields, see [1, Chapters 13-14]. In this course, we will primarily focus on the cases where $F = \mathbb{R}$, the field of real numbers, or $F = \mathbb{C}$, the field of complex numbers.

What is a vector? Mathematicians do not answer the question directly. Instead we answer the question what operations can be done on vectors. As long as these operations can be done, they are vectors.

Definition 1.1. A set V is called a *vector space* (*linear space*) over F if it is equipped with addition

$$V \times V \to V, \quad (x, y) \mapsto x + y$$

and scalar multiplication

$$F \times V \to V, \quad (a, x) \mapsto ax$$

satisfying the following axioms: for any $a, b \in F$, $x, y, z \in V$

(I) $x + y = y + x$, commutativity

(II) $(x + y) + z = x + (y + z)$, associativity

(III) There exists $0 \in V$ such that $x + 0 = x$, additive identity

(IV) For any $x \in V$, there exists $-x \in V$ such that $x + (-x) = 0$, additive inverse

(V) $a(bx) = (ab)x$, multiplicative associativity

(VI) $1x = x$, multiplicative identity

(VII) $a(x + y) = ax + ay$, distributive property 1

(VIII) $(a + b)x = ax + bx$. distributive property 2

An element $x \in V$ is called a *vector* and an element $a \in F$ is called a *scalar*.

Remark 1.2. The axioms above seems a lot, but they are just saying that the usual algebraic rules work for vector addition and scalar multiplication. This can be seen even more clearly in Proposition 1.3. It should be noted that the multiplication of two vectors is <u>not</u> defined within the structure of a vector space.

Proposition 1.3. *The following are direct consequences of the axioms*

(i) The element $0 \in V$ *is unique.*

(ii) *For any* $x \in V$*,* $0x = 0$*.*

(iii) *For any* $a \in F$*,* $a0 = 0$*.*

(iv) *If* $a \in F$*,* $x \in V$ *such that* $ax = 0$*, then either* $a = 0$ *(in* F*) or* $x = 0$ *(in* V*).*

(v) *For any* $x \in V$*, the element* $-x$ *is unique.*

(vi) *For any* $x \in V$*,* $(-1)x = -x$*.*

(vii) *If* $x + z = y + z$*, then* $x = y$*.*

Proof. (i)-(vi) are left as exercise. We now proof (vii). We have

$$x = x+0 = x+(z+(-z)) = (x+z)+(-z) = (y+z)+(-z) = y+(z+(-z)) = y+0 = y$$

where we used (III), (IV), (II), the assumption, (II), (IV), (III) in each equality. $\square$

Example 1.4.

$$F^n = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} : x_i \in F, i = 1, \ldots, n \right\}$$

equipped with addition

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

and scalar multiplication

$$c \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_n \end{pmatrix}.$$

The zero vector is $\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ and the additive inverse of $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ is $\begin{pmatrix} -x_1 \\ -x_2 \\ \vdots \\ -x_n \end{pmatrix}$.

Example 1.5. An $m \times n$ *matrix* A over F is an array of elements in F with m rows and n columns:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = (a_{ij})_{1 \le i \le m, 1 \le j \le n}$$

where $a_{ij} \in F$ for $1 \leq i \leq m$, $1 \leq j \leq n$. If the range of i, j are clear from context, we omit the range and write (a_{ij}). The set of all $m \times n$ matrices over F is denoted by $M_{m\times n}(F)$. It is equipped with the addition

$$(a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij})$$

and scalar multiplication

$$c(a_{ij}) = (ca_{ij}).$$

The zero vector is $0 = 0_{m\times n}$ an $m \times n$ matrix whose entries are all 0 and the additive inverse is$-(a_{ij}) = (-a_{ij})$. Then $M_{m\times n}(F)$ is a vector space. As a vector space, $M_{m\times n}(F)$ is simply $F^{m\times n}$, since a matrix can be flattened into a column vector with $m \times n$ components.

Given a matrix A its *transpose* is the matrix A^t given by transforming its rows into columns. For example

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}^t = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

It is a convention in math to write a vector in F^n as a column vector as we did in Example 1.4. But writing $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ in the notes all the time takes a lot of space, so for the rest of the notes we make use of the transpose notation and write $(x_1, \ldots, x_n)^t$.

Example 1.6. Let $k \in \mathbb{N}$ be a natural number and $C^k(\mathbb{R})$ be the set of functions $f : \mathbb{R} \to \mathbb{R}$ whose derivatives up to order k are continuous. We also write $C(\mathbb{R}) = C^0(\mathbb{R})$ for the set of continuous function. We equip $C^k(\mathbb{R})$ with the usual addition $(f+g)(x) = f(x)+g(x)$ and scalar multiplication $(cf)(x) = cf(x)$ for $c \in \mathbb{R}$. Then $C^k(\mathbb{R})$ is a vector space over $\mathbb{R}$.

Example 1.7. The space $\mathcal{P}_n(F)$ of polynomials of degree at most n consists of polynomials of the form

$$p(x) = a_n x^n + \cdots + a_1 x + a_0$$

where $a_i \in F$ for $0 \leq i \leq n$. Under the addition

$$(a_n x^n + \cdots + a_1 x + a_0) + (b_n x^n + \cdots + b_1 x + b_0) = (a_n + b_n)x^n + \cdots + (a_1 + b_1)x + (a_0 + b_0)$$

and scalar multiplication

$$c(a_n x^n + \cdots + a_1 x + a_0) = ca_n x^n + \cdots + ca_1 x + ca_0$$

$\mathcal{P}_n(F)$ is a vector space over F.

1.2 Subspaces

Definition 1.8. A subset $W \subset V$ of a vector space V is called a *subspace* if

(i) $0 \in W$

(ii) For any $x, y \in W$, we have $x + y \in W$

(iii) For any $a \in F$, $x \in W$, we have $ax \in W$

Lemma 1.9. *A subspace is a vector space.*

Proof. We first show that if $x \in W$ then $-x \in W$. By Proposition 1.3 (v) and definition of a subspace, $-x = (-1)x \in W$.

By the definition of a subspace, addition and scalar multiplication are well-defined on W. The axioms (I)-(VIII) are satisfied on W since they are satisfied for any vectors in V. $\square$

Example 1.10. A line through the origin in $\mathbb{R}^3$ is given by $\{tv : t \in \mathbb{R}\}$ where $v = (v_1, v_2, v_3)^t \neq 0$ is the direction of the line.

A plane through the origin in $\mathbb{R}^3$ is given by $\{tv + sw : t, s \in \mathbb{R}\}$ where v, w are two vectors in $\mathbb{R}^3$ not parallel to each other.

One can check that lines and planes are subspaces of $\mathbb{R}^3$. In fact $\{0\}$, lines, planes, $\mathbb{R}^3$ are the only subspaces of $\mathbb{R}^3$.

Example 1.11. A matrix A is called a *square* matrix if the number of columns is the same as the number of rows i.e. $A \in M_{n\times n}(F)$. A square matrix is *symmetric* if $A = A^t$. It is anti-symmetric if $A^t = -A$. The set of symmetric matrices and anti-symmetric matrices are subspaces of $M_{n\times n}(F)$ can be seen as follows: 0 matrix is both symmetric and anti-symmetric and for any $A, B \in M_{n\times n}(F)$ we have $(A + B)^t = A^t + B^t$ and $(cA)^t = cA^t$. Thus if A, B are symmetric, then $(A + B)^t = A^t + B^t = A + B$ and $(cA)^t = cA^t = cA$. The anti-symmetric case is similar.

Example 1.12. $C^k(\mathbb{R})$ is a subspace of $C^{k-1}(\mathbb{R})$ for any $k \in \mathbb{N}$. $\mathcal{P}_n(\mathbb{R})$ is a subspace of $C^k(\mathbb{R})$ for any $k \in \mathbb{N}$.

Solutions to a linear differential equation form a subspace.

Example 1.13. Let $V = \{f \in C^2(\mathbb{R}) : f'' + f = 0\}$ with the addition and scalar multiplication from $C^2(\mathbb{R})$. Then V is a subspace of $C^2(\mathbb{R})$ since $0 \in V$ and if $f, g \in V$ and $c \in \mathbb{R}$, we have $(f + g)'' + f + g = 0$ and $(cf)'' + cf = 0$. In fact, we know from ODE that $f(x) = c_1 \cos x + c_2 \sin x$ for $c_1, c_2 \in \mathbb{R}$ from which one can also check that V is a subspace. The point here is that we do not need the explicit formula for solutions to conclude that the set of all solutions form a subspace.

1.3 Span

We want to build the vector space out of the smallest amount of vectors. Addition and scalar multiplication are the only operations we have in a vector space. If we have $x, y \in V$, then $x + y$ and ax for $a \in F$ can be built out of them. Given a set of vectors $v_1, \ldots, v_n$ what are all vectors that can be built from them? This is the idea of linear combination and span.

Definition 1.14. Let V be a vector space over F. A *linear combination* of $v_1, \ldots, v_n \in V$ is a vector of the form

$$a_1 v_1 + \cdots + a_n v_n$$

where $a_1, \ldots, a_n \in F$. The set of all linear combinations of $v_1, \ldots, v_n$ is denoted by $\operatorname{span}(v_1, \ldots, v_n)$. We use the convention that $\operatorname{span} \emptyset = \{0\}$.

Proposition 1.15. $\operatorname{span}(v_1, \ldots, v_n)$ *is a subspace of* V*.*

Proof. By Proposition 1.3 (ii), we have $0 = 0v_1 + \cdots + 0v_n \in \operatorname{span}(v_1, \ldots, v_n)$. If $x = a_1 v_1 + \cdots + a_n v_n$ and $y = b_1 v_1 + \cdots + b_n v_n$ are two linear combinations, then $x + y = (a_1 + b_1)v_1 + \cdots + (a_n + b_n)v_n$ and $cx = (ca_1)v_1 + \cdots + (ca_n)v_n$ are both linear combinations of $v_1, \ldots, v_n$. Thus $\operatorname{span}(v_1, \ldots, v_n)$ is a subspace. $\square$

Example 1.16. Let $v_1 = (1, 0, 0)^t$. Then $\operatorname{span}(v_1) = \{(a_1, 0, 0)^t : a_1 \in \mathbb{R}\}$ is the line through the origin in the direction of v_1.

Let $v_2 = (0, 1, 0)^t$, then $\operatorname{span}(v_1, v_2) = \{(a_1, a_2, 0)^t : a_1, a_2 \in \mathbb{R}\}$ is the plane determined by 0, v and v_2.

Let $v_3 = (1, 1, 0)^t$, then $\operatorname{span}(v_1, v_2, v_3) = \{(a_1 + a_3, a_2 + a_3, 0)^t : a_1, a_2, a_3 \in \mathbb{R}\}$. This is the same as $\operatorname{span}(v_1, v_2)$ since it is just the set of points with $x_3 = 0$.

Let $v_4 = (0, 0, 1)^t$, then $\operatorname{span}(v_1, v_2, x) = \{(a_1, a_2, a_3)^t : a_1, a_2, a_3 \in \mathbb{R}\} = \mathbb{R}^3$.

Definition 1.17. $v_1, \ldots, v_n$ is a *spanning set* of V if $\operatorname{span}(v_1, \ldots, v_n) = V$.

If V has a finite spanning set, then V is called a *finite dimensional* vector space.

1.4 Linear dependence and independence

We see in Example 1.16 that v_3 is redundant in building vectors i.e. adding v_3 to v_1, v_2 does not change their span. We want to get rid of the redundant vectors to find the smallest set of vectors that span a vector space. The following concepts are important.

Definition 1.18. $v_1, \ldots, v_n$ is *linearly dependent* if there exist $a_1, \ldots, a_n \in F$ not all zero such that $a_1 v_1 + \cdots + a_n v_n = 0$.

$v_1, \ldots, v_n$ is *linearly independent* if or it is not linearly dependent, in other words, for any $a_1, \ldots, a_n \in F$ such that $a_1 v_1 + \cdots + a_n v_n = 0$ we have $a_1 = \cdots = a_n = 0$.

Example 1.19. $v_1 = (1, 0, 0)^t$, $v_2 = (0, 1, 0)^t$ are linearly independent since for any $a_1, a_2 \in F$ we have $0 = a_1 v_1 + a_2 v_2 = (a_1, a_2, 0)^t$ implies $a_1 = a_2 = 0$.

Example 1.20. $v_1 = (1, 0, 0)^t$, $v_2 = (0, 1, 0)^t$, $v_3 = (1, 1, 0)^t$ are linearly dependent since $v_1 + v_2 - v_3 = 0$.

The previous examples shows that a linearly dependent set of vectors is not small enough i.e. it has vectors that is redundant and that linearly independent vectors are not redundant. Thus our goal of finding a smallest set of vectors that builds the whole vector space is to find a basis:

Definition 1.21. $v_1, \ldots, v_n$ is a *basis* of V if $\text{span}(v_1, \ldots, v_n) = V$ and $v_1, \ldots, v_n$ are linearly independent.

References

[1] Dummit, David Steven, and Richard M. Foote. Abstract algebra. Vol. 3. Hoboken: Wiley, 2004.